

APPLICATION FOR UNITED STATES LETTERS PATENT

FOR

METHOD OF INVERSE QUANTIZING QUANTIZED SIGNAL SAMPLES OF AN IMAGE

DURING IMAGE DECOMPRESSION

Inventor(s): Tinku Acharya
Ping-Sing Tsai

Prepared by: Howard A. Skaist,
Senior Patent Attorney

Intel Corporation
2111 N.E. 25th Avenue; JF3-147
Hillsboro, OR 97124
Phone: (503) 264-0967
Facsimile: (503) 264-1729

"Express Mail" label number EL034434788US

**METHOD OF INVERSE QUANTIZING QUANTIZED SIGNAL SAMPLES
OF AN IMAGE DURING IMAGE DECOMPRESSION**

RELATED APPLICATION

ab
This patent application is related to concurrently filed U.S. patent application serial no. _____, titled "Method of Quantizing Signal Samples of an Image During Image Compression" by Acharya et al., (attorney docket no. 042390.P7135) assigned to the assignee of the current invention and herein incorporated by reference.

BACKGROUND

This disclosure relates to image compression and decompression and, more particularly, to inverse quantizing quantized signal samples of an image during image decompression.

As is well-known, in a variety of situations, it is desirable to have the ability to compress and decompress an image or a sequence of images. Without limitation, this may be desirable, for example, for transmission over a limited bandwidth communications channel, or for efficient storage of an image so it may later be retrieved. As is also well-known, image compression and decompression is typically computationally intensive and also employs a variety of operations in order to accomplish the end result, namely compression. For example, typically the image is transformed into another domain, such as from the spatial domain to the frequency

domain, and, likewise, typically some form of quantization is employed. Likewise, to decompress, a reverse process is employed. These are just two examples of the operations that may be performed upon the signal samples of an image in order to accomplish image compression and decompression. Additional operations may include entropy encoding, etc. Typically, each of these operations are computationally intensive. Therefore, if these operations are performed in software, a relatively large amount of processing power may be employed in order to perform the compression. Likewise, if, alternatively, these operations are performed using hardware, typically, separate hardware elements may be employed to perform each operation, to pipeline the processing, for example. A need, therefore, exists to reduce the amount of computationally intensive operations or the number of separate hardware elements employed to perform image compression and decompression.

BRIEF DESCRIPTION OF DRAWINGS

The subject matter regarded as the invention is particularly pointed out and distinctly claimed in the concluding portion of the specification. The invention, however, both as to organization, and method of operation, together with objects, features, and advantages thereof, may best be understood by reference to the following detailed description when read with the accompanying drawings in which:

Fig. 1 is a diagram schematically illustrating a process of applying an embodiment of a method for quantizing signal samples of an image during image

compression in accordance with the invention described in the aforementioned related patent application;

Fig. 2 is a schematic diagram illustrating an embodiment of a method of quantizing signal samples of an image during image compression in accordance with the invention described in the aforementioned related patent application in which a scaling operation is demonstrated;

Fig. 3 is a schematic diagram illustrating results of applying a Discrete Wavelet Transform (DWT) to an image where multiple or multi-levels are produced;

Fig. 4 is a schematic diagram illustrating results of applying an Inverse Discrete Wavelet Transform (IDWT) to an image where multiple or multi-levels are produced;

Fig. 5 is a block diagram illustrating a typical frame difference process employing a feedback loop in order to perform video compression; and

Fig. 6 is a schematic diagram illustrating an embodiment of a method of inverse quantizing quantized signal samples of an image during image decompression in accordance with the invention.

DETAILED DESCRIPTION

In the following detailed description, numerous specific details are set forth in order to provide a thorough understanding of the invention. However, it will be understood by those skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known methods, procedures, components, and circuits have not been described in detail so as not to obscure the present invention.

As previously indicated, one operation applied to accomplish compression of an image is transformation of an image from a first domain, such as a spatial domain, to a second domain, such as a frequency domain, although, of course, the invention is not limited in scope to transforming between these two particular domains. One common transformation employed in image compression is a Discrete Wavelet Transform (DWT) and also a multi-level DWT. Again, although the invention is not limited in scope in this respect, typically, for imaging compression or decompression, a two-dimensional DWT is employed. This is illustrated in more detail in Fig. 3.

Where the DWT is applied, the original image is decomposed into four subbands, illustrated, for example, by Fig 3(a). For a multi-level DWT, for example, this is referred to as a level of decomposition, and a subband at this level of decomposition may be further decomposed to produce another level of decomposition, illustrated, for example,

by Fig 3(b). In Fig 3, the LL subband of Fig 3(a), such as LL1, is further decomposed as shown in Fig 3(b).

As illustrated, for example, by Fig 3(a), the subbands are divided in the horizontal and the vertical direction, in this particular example according to frequency, with one half comprising the low frequency subband and the other half comprising the high frequency subband. Therefore, four different combinations of subbands result as illustrated, referred to as LL, LH, HL, and HH subbands, where L refers to low frequency signal information and H refers to high frequency signal information. It is generally understood that most of the basic signal information of the image is present in the low frequency LL subband, such as LL1. In a multi-level DWT, as previously indicated, the LL subband may be further decomposed, and this process may be taken to several levels, if desired. In this context, we shall designate the overall total number of levels of decomposition as K .

Typically, in image compression, once the DWT has been applied to the image and the so-called "transformed coefficients" have been obtained, then the coefficients are quantized based on subband location. It is desirable in a situation such as this to employ uniform scaling or quantization, meaning, in this context, that the transformed coefficients in a particular or selected subband are quantized based upon the same or substantially the same quantization value.

Quantization is desirable in image compression because it operates as another approach to filter the more useful information of the image from the less useful information of the image based, at least in part, on selected characteristics of a human vision system. Scalar quantization is performed by dividing a coefficient in the subband by a quantization or threshold value, Q . The nearest integer obtained after rounding or truncation, depending upon the implementation, for example, is the quantized coefficient. Tradeoffs or the selection of implementation details may, of course, be made for a variety of different reasons, such as accuracy or complexity, and, the invention is not limited in scope to a particular implementation detail, such as whether rounding or truncation is employed or how it may specifically be carried out.

As previously indicated, these operations are typically performed separately and may either result in a relatively high amount of computational complexity or, alternatively, may, in some cases, employ separate pieces of circuitry to perform the separate operations. Separating these operations in this manner, therefore, may result in lower performance in terms of speed and/or greater expense in terms of the amount or size of the silicon die where, for example, the operation is implemented through circuitry on an integrated circuit. It would, therefore, be desirable, for a variety of reasons, if the number of separate operations performed during image compression could be reduced.

Fig. 2 illustrates the application of an embodiment of a method of quantizing signal samples of an image during image compression in accordance with the invention

described in the aforementioned related patent application. In this particular embodiment, a process is applied to transform signal samples of the image from a first domain to a second domain such as, for example, as previously described, from the spatial domain to the frequency domain. However, in this particular embodiment, during the transformation process, signal samples are filtered by applying scaled or pre-scaled filter coefficients. Signal samples are filtered by applying scaled or pre-scaled filter coefficients along the image first in one direction such as, for example, row-wise, in this particular embodiment. This is illustrated by the left portion of Fig. 2. Therefore, in this particular embodiment, as rows of the image are decomposed into high and low frequency components, those components are also filtered by scaled or pre-scaled filter coefficients in accordance with the factors illustrated in Fig. 2, for this particular embodiment. In this embodiment, once the row-wise decomposition and filtering with scaled filter coefficients has been completed, it is then performed in a second direction, such as column-wise. Again, columns are decomposed into high and low frequency components and, likewise, the signal samples of those high and low frequency components are filtered with coefficients scaled in the manner illustrated by the right portion of Fig. 2. It will, of course, be appreciated that the invention described in the aforementioned related patent application is not limited in scope to this particular embodiment. For example, in other embodiments, alternatively decomposition and filtering with scaled or pre-scaled coefficients may be performed column-wise and then row-wise or in other alternative directions. Likewise, depending on the particular embodiment or approach, it may be that not all filter coefficients are scaled, for example. A variety of different approaches to scale the filter coefficients or portions

thereof, for example, may be employed. It is noted, however, that in this particular embodiment the two directions, horizontal and vertical, or row-wise and column-wise here, are mutually orthogonal, providing some benefit in terms of implementation because the two-dimensional scaling function may be expressed as a separable function. It shall be demonstrated hereinafter, when, for this particular embodiment, a transformation process of the image is complete, selected regions of the transformed signal samples, in this particular embodiment, the respective regions LL, HL, LH, and HH, are quantized by a common value.

In this particular embodiment, although the invention is not limited in scope in this respect, decomposition of an arbitrary sequence $x = \{x_0, x_1, \dots, x_{N-1}\}$ into a low pass subband $a = \{a_0, a_1, \dots, a_{(N/2)-1}\}$ and a high pass subband $c = \{c_0, c_1, \dots, c_{(N/2)-1}\}$ using 9-7 biorthogonal spline filters, for example, may be represented as follows:

$$a_n = \sum_{i=0}^8 l_i \cdot x_{2n-i} \quad [1]$$

$$c_n = \sum_{i=0}^6 h_i \cdot x_{2n-i} \quad [2]$$

where $l = \{l_0, l_1, \dots, l_8\}$ and $h = \{h_0, h_1, \dots, h_6\}$ are, respectively, the 9-tap low pass and the 7-tap high pass filter coefficients.

This is one technique for employing a two-dimensional DWT, although other approaches of implementing a two-dimensional DWT may be employed. For example, one may apply any 2-D filter as long as the filter may be decomposed into two 1-D filters. The expressions for the scaling factors remain the same. Likewise, a simple approach in this embodiment is to perform the decomposition row-wise and then column-wise, as previously described, although the invention is not limited in scope in this respect. As previously indicated, this is possible because, for this embodiment, the two-dimensional scaling function may be expressed as a separable function, that is, in this embodiment, the product of two one dimensional scaling functions.

It is likewise noted that, depending upon the particular embodiment, this method may be applied on multi-levels. For a K level two-dimensional DWT pyramid, where K is a positive integer, the filtering in the k^{th} level may be performed in the following manner:

First, at each level of DWT, set $Q(LL_k) = 1$ when level k is less than K .

ROW-WISE, ONE-DIMENSIONAL DWT

For low pass filtering over the LL_{k-1} subband or the input image when $k = 1$, each filter coefficient, l_i , is scaled by the factor:

$$\frac{1}{\sqrt{Q(LL_k)}}$$

For high pass filtering over the LL_{k-1} subband or the input image when $k = 1$, each filter coefficient, h_i , is scaled by the factor:

$$\frac{\sqrt{Q(LL_k)}}{Q(HL_k)}$$

COLUMN-WISE ONE-DIMENSIONAL DWT

For low pass filtering to generate subbands, LL_k and HL_k , each weight of the filter, l_i , is scaled by the factor:

$$\frac{1}{\sqrt{Q(LL_k)}}$$

For high pass filtering to generate the subband, LH_k , each weight of the filter, h_i , is scaled by the factor:

$$\frac{\sqrt{Q(LL_k)}}{Q(LH_k)}$$

For high pass filtering to generate subband HH_k , each weight of the filter, h_i , is scaled by the factor:

$$\frac{Q(HL_k)}{Q(HH_k)\sqrt{Q(LL_k)}}$$

where $Q(HL_k)$, $Q(HH_k)$, $Q(LH_k)$, and $Q(LL_k)$ are the quantization thresholds of the subbands, HL, HH, LH, and LL in the k^{th} level, respectively. Fig. 2 shows the corresponding scaling factors in terms of these subband locations, as previously described, where, for example, $Q(LL_k)$ for the k^{th} level is designated as Q_LL .

After mathematical manipulation, it may be demonstrated that after the row-wise and column-wise filtering of the four subbands at any level, the following scaling factors result:

$$\text{LL subband at level } k: \frac{1}{\sqrt{Q(LL_k)}} \frac{1}{\sqrt{Q(LL_k)}} = \frac{1}{Q(LL_k)} \text{ when } k = K, \text{ or } 1 \text{ when } k < K;$$

$$\text{HL subband at level } k: \frac{\sqrt{Q(LL_k)}}{Q(HL_k)} \frac{1}{\sqrt{Q(LL_k)}} = \frac{1}{Q(HL_k)};$$

$$\text{LH subband at level } k: \frac{1}{\sqrt{Q(LL_k)}} \frac{\sqrt{Q(LL_k)}}{Q(LH_k)} = \frac{1}{Q(LH_k)};$$

$$\text{HH subband at level } k: \frac{\sqrt{Q(LL_k)}}{Q(HL_k)} \frac{Q(HL_k)}{Q(HH_k)\sqrt{Q(LL_k)}} = \frac{1}{Q(HH_k)}.$$

Therefore, the desired scaling or quantization in designated or selected regions has been accomplished. Fig. 1 is an example illustrating a process where this embodiment is applied for $K = 3$.

As previously demonstrated, for this particular embodiment of a method of quantizing signal samples of an image during image compression, uniform scaling or

quantization has been integrated into a multi-level DWT transformation. As a result, quantization circuitry may be eliminated. This may, therefore, result in a smaller silicon die for an integrated circuit, potentially reducing cost and improving speed. A similar approach may be applied for a software implementation in order to reduce computational "costs" and achieve faster execution of a DWT based image compression technique, for example.

Fig. 5 is a block diagram illustrating a typical frame difference process with a feedback loop in order to perform video compression. As illustrated by the block diagram in the forward path, a DWT operation and a quantization operation are respectively applied by blocks 510 and 520. This corresponds to the approach normally taken to perform compression, as previously described. Therefore, a method of quantizing signal samples of an image during image compression in accordance with the invention described in the aforementioned related patent application, as previously described, may be employed. Nonetheless, Fig. 5 further illustrates the application of an inverse quantization operation 530 and an inverse DWT 540 (IDWT). It would be desirable to have the ability to combine both the inverse quantization operation and the inverse DWT operation, such as the one provided in the feedback loop illustrated by 535 for the inverse DWT and inverse quantization operations, for example. In addition to the advantages previously described of applying the previously described embodiment to the forward loop in place of 510 and 520 by combining them, advantages are also gained by similarly combining 530 and 540. This would further simplify and reduce the size of the circuit implementing the compression operation

illustrated by Fig. 5 and, likewise, where the implementation is performed in software, a reduction in computation time may result.

In this particular embodiment of a method of inverse quantizing quantized signal samples of an image during image decompression in accordance with the present invention, a process to transform the quantized signal samples from the first domain, such as the frequency domain, to a second domain, such as the spatial domain is applied. During the transformation process, quantized signal samples are filtered by applying scaled or pre-scaled filter coefficients, in this particular embodiment. The signal samples are first filter along the image in a first direction by applying scaled or pre-scaled filter coefficients, such as column-wise, and then the signal samples are filtered along the image in a second direction such as row-wise, by applying scaled or pre-scaled filter coefficients, so that at the completion of the transformation process of the image, the transformed signal samples are inverse quantized. It is noted, as with the previously described embodiment of a method of quantizing signal samples of an image during image compression, the first direction and second direction are mutually orthogonal, in this embodiment. Likewise, although the invention is not limited in scope in this respect, the transformation process may comprise the inverse DWT. Furthermore , the invention is not restricted in scope to a particular direction along the image that is filtered or to a particular order in which these directions are filterd by applying scaled coefficients. For example, alternatively, in another embodiment, it may be possible to filter signal samples with scaled coefficients first row-wise and then column-wise.

The following embodiment is described in terms of a two-dimensional inverse DWT, although, again, the invention is not limited in scope in this respect. Likewise, this particular embodiment may be applied to a K level pyramid where K is a positive integer, although, again, the invention is not limited in the scope to being applied to more than one level, or to a particular number of levels. Nonetheless, for this embodiment, inverse filtering the k^{th} level may be performed in the following manner.

First, at a level of the transformation process, in this particular embodiment, an IDWT process is applied, and $Q(LL_k) = 1$ when k is less than K . As previously discussed, as the transformation process is applied, transformed signal samples filtered along the image in a first direction, such as column-wise, by applying scaled filter coefficients. Therefore, where, for this particular embodiment, an IDWT is being employed, scaling of the filter coefficients is performed as follows:

For inverse low pass filtering over the DWT coefficients at level k , subbands LL_k and HL_k , filter coefficients are scaled by the factor $\sqrt{Q(LL_k)}$.

Next, for inverse high pass filtering over the DWT coefficients at level k , subband LH_k , filter coefficients are scaled by the factor $\frac{Q(LH_k)}{\sqrt{Q(LL_k)}}$.

Then, for inverse high pass filtering over the DWT coefficients at level k , subband

HH_k , filter coefficients are scaled by the factor $\frac{Q(HH_k)\sqrt{Q(LL_k)}}{Q(HL_k)}$.

As previously described, after filtering is performed along the image in the first direction, then filtering is performed along the image in a second direction, in this particular embodiment row-wise, as follows:

For inverse low pass filtering over the DWT coefficients at level k , subbands LL_k and LH_k , filter coefficients are scaled by the factor $\sqrt{Q(LL_k)}$.

For inverse high pass filtering over the DWT coefficients at level k , subband HL_k and HH_k , filter coefficients are scaled by the factor $\frac{Q(HL_k)}{\sqrt{Q(LL_k)}}$,

where $Q(LH_k)$, $Q(HH_k)$, $Q(LH_k)$ and $Q(LL_k)$ are the quantization thresholds for the subbands HL , HH , LH , and LL in the k^{th} level, respectively.

Fig. 6 shows corresponding scale factors for filter coefficients in terms of subband locations. It is noted that these are the multiplicative inverses of the scale factors applied to filter coefficients in the particular embodiment of the invention described in the aforementioned related patent application. It may be demonstrated by mathematical manipulation that after the column-wise and row-wise inverse filtering of the DWT coefficients at any level, scaling by the following factors occurs.

LL sub-band at level k : $\sqrt{Q(LL_k)}\sqrt{Q(LL_k)} = Q(LL_k)$ when $k = K$, or 1 when $k < K$;

HL sub-band at level k : $\sqrt{Q(LL_k)} \frac{Q(HL_k)}{\sqrt{Q(LL_k)}} = Q(HL_k)$;

LH sub-band at level k : $\frac{Q(LH_k)}{\sqrt{Q(LL_k)}} \sqrt{Q(LL_k)} = Q(LH_k)$;

HH sub-band at level k : $\frac{Q(HH_k)\sqrt{Q(LL_k)}}{Q(HL_k)} \frac{Q(HL_k)}{\sqrt{Q(LL_k)}} = Q(HH_k)$.

Fig. 4 shows an example with $K = 3$. Again, it is noted that this example shows multilevel inverses of previously discussed scale factors. Hence, for this embodiment, the desired inverse uniform scale quantization has resulted. As previously described, as a result, the inverse quantization circuitry may be eliminated, potentially resulting in a smaller silicon die for large scale integration (LSI) implementation and potentially an improvement in speed performance. Likewise, implementation in software may reduce the computation cost and faster execution of the DWT based image compression/decompression technique.

As previously indicated, it will be appreciated that many different embodiments are within the scope of the present invention. For example, although a method of inverse quantizing quantized signal samples of an image during image decompression in accordance with the present invention is described, alternatively, an embodiment may comprise, for example, a device, such as, for example, an integrated circuit. In such an embodiment, the integrated circuit may, for example, include input ports to receive signal samples associated with at least one image and the integrated circuit may

include digital circuitry, although, of course, the invention is not limited in scope in this respect. The digital circuitry may have a configuration to apply a process to transform the signal samples from a first domain to a second domain and, during the transform process, filter signal samples, by first applying scaled filter coefficients to signal samples along the image in a first direction and then applying scaled filter coefficients to signal samples along the image in a second direction, so that, at the completion of the transform process of the image, at least selected regions of the transformed signal samples are inverse quantized. Likewise, other aspects of the previously described embodiment, for example, may be included, although, again, the invention is not limited in scope in this respect. Additionally, features not previously described may be included in an embodiment in accordance with the invention.

Likewise, yet another an alternative embodiment may comprise an article, such as a storage medium, having stored thereon, instructions, that, when executed by a system to execute such instructions, for example, may result in the previously described method embodiment being performed. For example, although the invention is not limited in scope in this respect, the system may comprise a computer, such as a personal computer, a set-top box, a server, or anyone of a number of other systems that may be adapted to execute specific instructions stored on a storage medium.

While certain features of the invention have been illustrated and described herein, many modifications, substitutions, changes and equivalents will now occur to those skilled in the art. It is, therefore, to be understood that the appended claims are

